Demo Abstract: BiGuide: A Bi-level Data Acquisition Guidance for Object Detection on Mobile Devices

Lin Duan
Duke University
USA
lin.duan@duke.edu

Ying Chen
Duke University
USA
ying.chen151@duke.edu

Maria Gorlatova
Duke University
USA
maria.gorlatova@duke.edu

ABSTRACT

Real-time object detection (OD) is a key enabling technology for a wide range of emerging mobile system applications. However, deploying an OD model pre-trained on a public dataset (*source domain*) in a specific local environment (*target domain*) is known to lead to significant performance degradation because of the so-called *domain gap* between the dataset and the environment. Collecting local data and fine-tuning the OD model on this data is a commonly used approach for improving the robustness of OD models in real-world deployments. Yet, the question of *how* to collect this data is currently largely overlooked; unsupported data collection is likely to produce datasets that contain significant proportion of redundant or uninformative data for model training. In this demo, we present BiGuide, a bi-level image data acquisition guidance for OD tasks, to guide users to change their camera locations or angles to different extents (significantly or slightly) to obtain the data which benefits model training via image-level and object instance-level guidance. We showcase an interactive demonstration of collecting data for a lemur species detection application we are developing and deploying at the Duke Lemur Center. Demo participants will take pictures of lemur toys with the mobile phone under the real-time guidance and will observe the real-time display of the metrics that assess the importance of the captured data. They will develop an intuition for how real-time image importance assessment and bi-level guidance improve the quality of collected data.

CCS CONCEPTS

• Computing methodologies → Machine learning; **• Human-centered computing → Ubiquitous and mobile computing systems and tools**.

KEYWORDS

Data acquisition, user guidance, robust object detection, domain adaptation, edge computing, uncertainty and diversity evaluation.

ACM Reference Format:
Lin Duan, Ying Chen, and Maria Gorlatova. 2023. Demo Abstract: BiGuide: A Bi-level Data Acquisition Guidance for Object Detection on Mobile Devices. In *The 22nd International Conference on Information Processing in Sensor Networks (IPSN '23), May 9–12, 2023, San Antonio, TX, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3583120.3589809

IPSN '23, May 9–12, 2023, San Antonio, TX, USA

ACM ISBN 979-8-4007-0118-4/23/05…$15.00
https://doi.org/10.1145/3583120.3589809

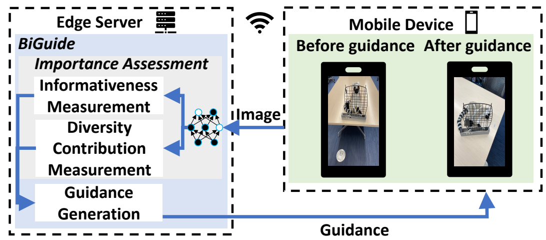

Figure 1: BiGuide system overview. The mobile device streams images to the edge. The edge conducts importance assessment, and generates and sends back the guidance.

1 INTRODUCTION

Object detection (OD) plays a crucial role in a wide spectrum of applications, including augmented reality and autonomous driving. However, when applying a well-trained OD model to local environments in the wild, the *domain gap* between training samples (source domain) and the local environments (target domain) degrades the detection accuracy. Collecting local image data to fine-tune the model is crucial for adapting the OD model to the local environment. While researchers usually presume the availability of the local labeled data for model training [4], *how* such data is obtained is currently largely overlooked. Data acquisition without user guidance lacks a measure of data quality and tends to include redundant and uninformative data that degrades the model performance.

Current data collection methods to improve the performance of different applications mainly focus on maximizing the coverage of the viewpoints. For the spatial mapping and environment monitoring tasks, some works [3] optimally place and steer robots' cameras to collect useful data. Yet, these works require an exhaustive search for optimal locations and angles from known alternatives. For OD tasks, LabelAR [2] guides users to move the camera to cover a wide variety of viewpoints for the object of interest. But this method cannot guide the data acquisition process without pre-set viewpoints. Furthermore, not all the data that covers a wide range of viewpoints contributes to model training. Therefore, effective real-time data acquisition guidance is needed to effectively guide users to collect data that are useful for model training.

In this demo, we present BiGuide, an edge computing-based bi-level data acquisition guidance with the context-awareness capability that evaluates the importance of the image data in the local environment for OD model training. As *informative* and *diverse* data contributes significantly to model training [1], BiGuide instructs users to collect informative and diverse data via image-level and object instance-level guidance in real time. We define uncertainty to quantify the informativeness of the current data; we assess how much the current image contributes to the diversity of the Source domain and Previously collected images (SP images) at both image and object instance levels by analyzing the dissimilarities between the current image and SP images. Based on that, BiGuide generates bi-level guidance shown on mobile devices for users. *BiGuide is the first tool that instructs users to collect diverse and informative data for training OD models.* We showcase that BiGuide can guide users to collect high-quality data for the lemur species detection application we are developing to help study and manage the lemurs (e.g., in cages) at the Duke Lemur Center, which has a large domain gap with the source domain (e.g., with unobstructed environments).

2 SYSTEM DESIGN

The system overview of BiGuide is shown in Figure 1. We implement an edge computing-based architecture to generate the data acquisition guidance. The mobile app running on the mobile device streams images to the edge in real time. The edge generates the data acquisition guidance and sends it wirelessly to the mobile device.

Importance assessment: As the server receives the image, BiGuide evaluates the importance of the image through *informativeness and diversity contribution* measurements. We adopt the well-known uncertainty evaluation approach [1] to measure the informativeness of the current image. The uncertainty is determined by the prediction probability of the OD model. Data with high uncertainty means the OD model is uncertain about the prediction results of the data, which indicates that the data is informative for the model training. To complement the importance assessment, inspired by [1] that uses the diversity to measure the importance of the data, we propose the *image-level and object instance-level diversity scores* of the new image data to quantify how much the current image contributes to the diversity of the SP images at both image and instance levels. The diversity score is calculated by measuring the feature distance (calculated by the OD model) between the current image and the cluster centers of SP images. A high diversity score indicates that the captured image differs from SP images and tends to extend the knowledge of images learned by the OD model.

Guidance generation: BiGuide generates image-level and object instance-level guidance to help users collect informative and diverse data. When the image-level diversity score is low, image-level guidance instructs users to slightly move their camera locations to get different image backgrounds. When the instance-level diversity score is low, instance-level guidance instructs users to slightly change their camera orientations or wait for the object's pose change to get more diverse instances. In addition, when the uncertainty is low, the guidance instructs users to move their camera locations to a significantly different viewpoint. Otherwise, feedback is sent to users to acknowledge the image's high quality.

3 INTERACTIVE DEMONSTRATION

We showcase an application of BiGuide for lemur species detection we are developing and deploying at the Duke Lemur Center. The demonstration follows the workflow as in Figure 1. Demo participants can gain an understanding of the advantages of using BiGuide to collect important data for model training, and have the opportunity to directly observe the informative and diverse nature of the data they collect. A video of the demo is available online.[1]

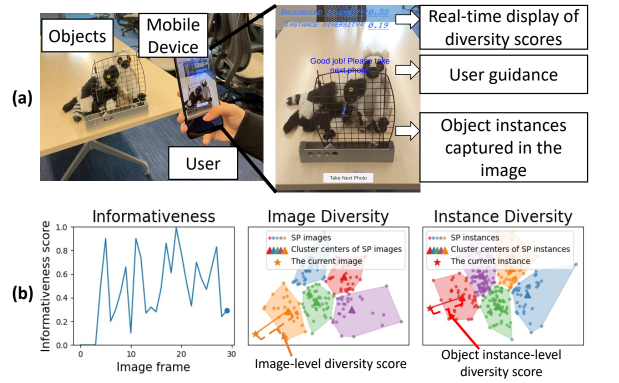

Figure 2: BiGuide in action: (a) The user points the camera towards the lemurs and observes real-time guidance. (b) Real-time visualization of informativeness and diversity scores.

The demo is performed using a Google Pixel 3 XL mobile phone running Android 11. A Lenovo laptop with an AMD Ryzen 4700H CPU and an NVIDIA GTX 1660 Ti GPU serves as the edge server. The mobile app is built with Unity 2020.3.14f and ARCore 4.1.7.

Our portable demo setup consists of several lemur toys of different species and cage doors. To mimic the caged lemurs at the Duke Lemur Center, which has a domain gap with the source domain (e.g., with unobstructed environments), we place lemur toys on the table behind the cage doors, similar to the illustration in Figure 2. Participants play the role of the wildlife center's curators and help collect data to train a robust OD model for studying and managing the lemurs. They interact with the demo by holding the mobile phone and pointing the camera toward the lemurs. The participants see the real-time data acquisition guidance sent by the server and displayed on the mobile phone. According to the guidance, they can either take the image as is, or significantly or slightly adjust their locations, camera angles, or lemur toys' poses. At the same time, they can observe how informative and diverse the image they collect is directly from the real-time visualization on the screen of the laptop (i.e., the server). Demo participants can understand how real-time image importance assessment and bi-level guidance improve the quality of collected data and aid in model training.

ACKNOWLEDGMENTS

This work was supported in part by NSF grants CSR-1903136 and CNS-1908051, NSF CAREER Award IIS-2046072, and a Facebook Research Award.

REFERENCES

[1] Andreas Kirsch, Joost Van Amersfoort, and Yarin Gal. 2019. BatchBALD: Efficient and diverse batch acquisition for deep bayesian active learning. In *Proc. NeurIPS.*
[2] Michael Laielli, James Smith, Giscard Biamby, Trevor Darrell, and Bjoern Hartmann. 2019. LabelAR: A spatial guidance interface for fast computer vision image collection. In *Proc. ACM UIST.*
[3] Teng Li, Chaoqun Wang, Max Q.-H. Meng, and Clarence W. de Silva. 2022. Attention-driven active sensing with hybrid neural network for environmental field mapping. *IEEE Trans. Autom. Sci. Eng.* 19, 3 (2022), 2135–2152.
[4] Xian Shuai, Yulin Shen, Siyang Jiang, Zhihe Zhao, Zhenyu Yan, and Guoliang Xing. 2022. BalanceFL: Addressing class imbalance in long-tail federated learning. In *Proc. ACM/IEEE IPSN.*

[1]Link to the demo video: https://sites.duke.edu/linduan/